

SEQUENCE LISTING

<110> Lanes, Olav
Willasen, Nils Peder
Guddal, Per Henrik
Gjellesvik, Dag Rune

<120> Cod uracil-DNA glycosylase, gene coding therefore,
recombinant DNA containing said gene or operative parts
thereof, a method for preparing said protein and the
use of said protein or said operative pa

<130> U013209-3

<140> 09/758,017
<141> 2001-01-10

<150> 2000 5428
<151> 2000-10-27

<150> 2000 0163
<151> 2000-01-12

<160> 19

<170> PatentIn Ver. 2.0

<210> 1
<211> 1283
<212> DNA
<213> Gadus morhua

<220>
<221> CDS
<222> (18)...(920)

<400> 1
gacatccgct tgcaa atg ttg ttc aag tta ggg tta tgc caa aga tgc 50
 Met Leu Phe Lys Leu Gly Leu Cys Gln Arg Cys
 1 5 10

ata tca tca aat cgg gtg tta cca ggt tta cta att ccc caa act tta 98
Ile Ser Ser Asn Arg Val Leu Pro Gly Leu Leu Ile Pro Gln Thr Leu
 15 20 25

tgt ttt tct aaa tta atg aag ata acg ccg aag aaa ctg agg tcc tca 146
Cys Phe Ser Lys Leu Met Lys Ile Thr Pro Lys Lys Leu Arg Ser Ser
 30 35 40

| | | | |
|--|-----|-----|-----|
| aat gtg gaa caa aag acg tca tcg cca cag ctt tca gtg gag cag ctg | | | 194 |
| Asn Val Glu Gln Lys Thr Ser Ser Pro Gln Leu Ser Val Glu Gln Leu | | | |
| 45 | 50 | 55 | |
| gaa aga atg gcc aaa aat aag aaa gca gcg ctt gac aag att aga gca | | | 242 |
| Glu Arg Met Ala Lys Asn Lys Ala Ala Leu Asp Lys Ile Arg Ala | | | |
| 60 | 65 | 70 | 75 |
| aaa gca acg cct gca ggt ttc gga gag act tgg aga aga gag ctg gct | | | 290 |
| Lys Ala Thr Pro Ala Gly Phe Gly Glu Thr Trp Arg Arg Glu Leu Ala | | | |
| 80 | 85 | 90 | |
| gca gag ttt gaa aag cca tac ttc aaa caa ttg atg tcc ttt gta gct | | | 338 |
| Ala Glu Phe Glu Lys Pro Tyr Phe Lys Gln Leu Met Ser Phe Val Ala | | | |
| 95 | 100 | 105 | |
| gat gag agg agc cgt cac acc gtc tac cca ccg gct gat caa gtg tac | | | 386 |
| Asp Glu Arg Ser Arg His Thr Val Tyr Pro Pro Ala Asp Gln Val Tyr | | | |
| 110 | 115 | 120 | |
| agt tcg aca gag atg tgt gac att caa gat gtg aaa gta gtg att cta | | | 434 |
| Ser Ser Thr Glu Met Cys Asp Ile Gln Asp Val Lys Val Val Ile Leu | | | |
| 125 | 130 | 135 | |
| ggc cag gac cct tac cac ggt ccc aac caa gca cat gga ctc tgt ttc | | | 482 |
| Gly Gln Asp Pro Tyr His Gly Pro Asn Gln Ala His Gly Leu Cys Phe | | | |
| 140 | 145 | 150 | 155 |
| agt gtg caa aag cca gtt ccc cct ccc agt ctc gtg aac ata tac | | | 530 |
| Ser Val Gln Lys Pro Val Pro Pro Ser Leu Val Asn Ile Tyr | | | |
| 160 | 165 | 170 | |
| aaa gaa ttg tgt acc gac att gat ggc ttc aag cat cct gga cat gga | | | 578 |
| Lys Glu Leu Cys Thr Asp Ile Asp Gly Phe Lys His Pro Gly His Gly | | | |
| 175 | 180 | 185 | |
| gat cta agc gga tgg gca aaa caa ggg gtg ctg ctg ctt aac gcg gtg | | | 626 |
| Asp Leu Ser Gly Trp Ala Lys Gln Gly Val Leu Leu Leu Asn Ala Val | | | |
| 190 | 195 | 200 | |
| ctg acc gtg cggt gcc cat cag gcc aac tcc cac aag gac aga ggc tgg | | | 674 |
| Leu Thr Val Arg Ala His Gln Ala Asn Ser His Lys Asp Arg Gly Trp | | | |
| 205 | 210 | 215 | |
| gag acc ttc acc gac gct gtg atc aag tgg ctg agc gtc aac cggt gaa | | | 722 |
| Glu Thr Phe Thr Asp Ala Val Ile Lys Trp Leu Ser Val Asn Arg Glu | | | |
| 220 | 225 | 230 | 235 |

| | | |
|---|------|------|
| gga gtc gtt ttc ctg ttg tgg ggc tca tac gcc cat aag aag gga gcg | | 770 |
| Gly Val Val Phe Leu Leu Trp Gly Ser Tyr Ala His Lys Lys Gly Ala | | |
| 240 | 245 | 250 |
| acc atc gac agg aaa cgt cac cat gtc ttg caa gct gtt cat cca tct | | 818 |
| Thr Ile Asp Arg Lys Arg His His Val Leu Gln Ala Val His Pro Ser | | |
| 255 | 260 | 265 |
| cct ttg tct gct cat cgt ggg ttc ctt ggt tgt aag cac ttc tcc aag | | 866 |
| Pro Leu Ser Ala His Arg Gly Phe Leu Gly Cys Lys His Phe Ser Lys | | |
| 270 | 275 | 280 |
| gct aac ggg ctg ctg aaa cta tct ggg acg gag cct ata aac tgg aga | | 914 |
| Ala Asn Gly Leu Leu Lys Leu Ser Gly Thr Glu Pro Ile Asn Trp Arg | | |
| 285 | 290 | 295 |
| gca ctc taactcttta tgctgcctta tactgttaac tgtttaaga tgaacatcac | | 970 |
| Ala Leu | | |
| 300 | | |
| actatatattt ctacagcttt tccaaaggta aaccaatcta taagctttca tttgtcttt | 1030 | |
| ggaatgatgc tgctttggc cggttttaga tacttaaaac actttaccac tctgccatgt | 1090 | |
| tgactcatgt tcagtcaata taactttcac aacttgaaca aaaatgttat tttataattg | 1150 | |
| attatatattct gtacattaaa gattgtttt ttcccaggct gtttcataagg tactaggata | 1210 | |
| ttaaactgtt attaacctat tttccatgat gtcaactgct taagtttta tgcagaaata | 1270 | |
| aattatataat tta | | 1283 |

<210> 2

<211> 301

<212> PRT

<213> Gadus morhua

<400> 2

| | | | |
|---|---|----|----|
| Met Leu Phe Lys Leu Gly Leu Cys Gln Arg Cys Ile Ser Ser Asn Arg | | | |
| 1 | 5 | 10 | 15 |

| | | | |
|---|----|----|--|
| Val Leu Pro Gly Leu Leu Ile Pro Gln Thr Leu Cys Phe Ser Lys Leu | | | |
| 20 | 25 | 30 | |

| | | | |
|---|----|----|--|
| Met Lys Ile Thr Pro Lys Lys Leu Arg Ser Ser Asn Val Glu Gln Lys | | | |
| 35 | 40 | 45 | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Thr | Ser | Ser | Pro | Gln | Leu | Ser | Val | Glu | Gln | Leu | Glu | Arg | Met | Ala | Lys |
| 50 | | | | | | | | | | | | | | | 60 |
| Asn | Lys | Lys | Ala | Ala | Leu | Asp | Lys | Ile | Arg | Ala | Lys | Ala | Thr | Pro | Ala |
| 65 | | | | | | | | | | | | | | | 80 |
| Gly | Phe | Gly | Glu | Thr | Trp | Arg | Arg | Glu | Leu | Ala | Ala | Glu | Phe | Glu | Lys |
| | | | | | | | | | | | | | | | 95 |
| 85 | | | | | | | | | | | | | | | |
| Pro | Tyr | Phe | Lys | Gln | Leu | Met | Ser | Phe | Val | Ala | Asp | Glu | Arg | Ser | Arg |
| | | | | | | | | | | | | | | | 100 |
| | | | | | | | | | | | | | | | 105 |
| | | | | | | | | | | | | | | | 110 |
| His | Thr | Val | Tyr | Pro | Pro | Ala | Asp | Gln | Val | Tyr | Ser | Ser | Thr | Glu | Met |
| | | | | | | | | | | | | | | | 115 |
| | | | | | | | | | | | | | | | 120 |
| | | | | | | | | | | | | | | | 125 |
| Cys | Asp | Ile | Gln | Asp | Val | Lys | Val | Val | Ile | Leu | Gly | Gln | Asp | Pro | Tyr |
| | | | | | | | | | | | | | | | 130 |
| | | | | | | | | | | | | | | | 135 |
| | | | | | | | | | | | | | | | 140 |
| His | Gly | Pro | Asn | Gln | Ala | His | Gly | Leu | Cys | Phe | Ser | Val | Gln | Lys | Pro |
| | | | | | | | | | | | | | | | 145 |
| | | | | | | | | | | | | | | | 150 |
| | | | | | | | | | | | | | | | 155 |
| | | | | | | | | | | | | | | | 160 |
| Val | Pro | Pro | Pro | Pro | Ser | Leu | Val | Asn | Ile | Tyr | Lys | Glu | Leu | Cys | Thr |
| | | | | | | | | | | | | | | | 165 |
| | | | | | | | | | | | | | | | 170 |
| | | | | | | | | | | | | | | | 175 |
| Asp | Ile | Asp | Gly | Phe | Lys | His | Pro | Gly | His | Gly | Asp | Leu | Ser | Gly | Trp |
| | | | | | | | | | | | | | | | 180 |
| | | | | | | | | | | | | | | | 185 |
| | | | | | | | | | | | | | | | 190 |
| Ala | Lys | Gln | Gly | Val | Leu | Leu | Leu | Asn | Ala | Val | Leu | Thr | Val | Arg | Ala |
| | | | | | | | | | | | | | | | 195 |
| | | | | | | | | | | | | | | | 200 |
| | | | | | | | | | | | | | | | 205 |
| His | Gln | Ala | Asn | Ser | His | Lys | Asp | Arg | Gly | Trp | Glu | Thr | Phe | Thr | Asp |
| | | | | | | | | | | | | | | | 210 |
| | | | | | | | | | | | | | | | 215 |
| | | | | | | | | | | | | | | | 220 |
| Ala | Val | Ile | Lys | Trp | Leu | Ser | Val | Asn | Arg | Glu | Gly | Val | Val | Phe | Leu |
| | | | | | | | | | | | | | | | 225 |
| | | | | | | | | | | | | | | | 230 |
| | | | | | | | | | | | | | | | 235 |
| | | | | | | | | | | | | | | | 240 |
| Leu | Trp | Gly | Ser | Tyr | Ala | His | Lys | Lys | Gly | Ala | Thr | Ile | Asp | Arg | Lys |
| | | | | | | | | | | | | | | | 245 |
| | | | | | | | | | | | | | | | 250 |
| | | | | | | | | | | | | | | | 255 |
| Arg | His | His | Val | Leu | Gln | Ala | Val | His | Pro | Ser | Pro | Leu | Ser | Ala | His |
| | | | | | | | | | | | | | | | 260 |
| | | | | | | | | | | | | | | | 265 |
| | | | | | | | | | | | | | | | 270 |
| Arg | Gly | Phe | Leu | Gly | Cys | Lys | His | Phe | Ser | Lys | Ala | Asn | Gly | Leu | Leu |
| | | | | | | | | | | | | | | | 275 |
| | | | | | | | | | | | | | | | 280 |
| | | | | | | | | | | | | | | | 285 |
| Lys | Leu | Ser | Gly | Thr | Glu | Pro | Ile | Asn | Trp | Arg | Ala | Leu | | | |
| | | | | | | | | | | | | | | | 290 |
| | | | | | | | | | | | | | | | 295 |
| | | | | | | | | | | | | | | | 300 |

<210> 3
 <211> 1355
 <212> DNA
 <213> Gadus morhua

<220>
 <221> CDS
 <222> (90)..(992)

<400> 3
 gatggtttag gaggatagta ctttgacact gtttagcgaa gggaaaaacg gagttattgt 60
 gcatatcgtt ttagccctac gttaaaaa atg att ggt caa cag cat atc aac 113
 Met Ile Gly Gln Gln His Ile Asn
 1 5

tct ttc ttc tca cca gtt tca aaa aag aga gtt tca aag gaa tta ggt 161
 Ser Phe Phe Ser Pro Val Ser Lys Lys Arg Val Ser Lys Glu Leu Gly
 10 15 20

aaa acc gaa aag cat gcc gaa gaa gtt cag ata acg ccg aag aaa ctg 209
 Lys Thr Glu Lys His Ala Glu Glu Val Gln Ile Thr Pro Lys Lys Leu
 25 30 35 40

agg tcc tca aat gtg gaa caa aag acg tca tcg cca cag ctt tca gtg 257
 Arg Ser Ser Asn Val Glu Gln Lys Thr Ser Ser Pro Gln Leu Ser Val
 45 50 55

gag cag ctg gaa aga atg gcc aaa aat aag aaa gca gcg ctt gac aag 305
 Glu Gln Leu Glu Arg Met Ala Lys Asn Lys Lys Ala Ala Leu Asp Lys
 60 65 70

att aga gca aaa gca acg cct gca ggt ttc gga gag act tgg aga aga 353
 Ile Arg Ala Lys Ala Thr Pro Ala Gly Phe Gly Glu Thr Trp Arg Arg
 75 80 85

gag ctg gct gca gag ttt gaa aag cca tac ttc aaa caa ttg atg tcc 401
 Glu Leu Ala Ala Glu Phe Glu Lys Pro Tyr Phe Lys Gln Leu Met Ser
 90 95 100

ttt gta gct gat gag agg agc cgt cac acc gtc tac cca ccg gct gat 449
 Phe Val Ala Asp Glu Arg Ser Arg His Thr Val Tyr Pro Pro Ala Asp
 105 110 115 120

caa gtg tac agt tgg aca gag atg tgt gac att caa gat gtg aaa gta 497
 Gln Val Tyr Ser Trp Thr Glu Met Cys Asp Ile Gln Asp Val Lys Val

| 125 | 130 | 135 | |
|---|-----|-----|------|
| gtg att cta ggc cag gac cct tac cac ggt ccc aac caa gca cat gga Val Ile Leu Gly Gln Asp Pro Tyr His Gly Pro Asn Gln Ala His Gly 140 | 145 | 150 | 545 |
| ctc tgt ttc agt gtg caa aag cca gtt ccc cct ccc ccc agt ctc gtg Leu Cys Phe Ser Val Gln Lys Pro Val Pro Pro Pro Ser Leu Val 155 | 160 | 165 | 593 |
| aac ata tac aaa gaa ttg tgt acc gac att gat ggc ttc aag cat cct Asn Ile Tyr Lys Glu Leu Cys Thr Asp Ile Asp Gly Phe Lys His Pro 170 | 175 | 180 | 641 |
| gga cat gga gat cta agc gga tgg gca aac aag ggg tgc tgc tgc tta Gly His Asp Leu Ser Gly Trp Ala Asn Lys Gly Cys Cys Cys Leu 185 | 190 | 195 | 200 |
| acg cgc tgc ctg acc gtg cg ^g gcc cat cag gcc aac tcc cac aag gac Thr Arg Cys Leu Thr Val Arg Ala His Gln Ala Asn Ser His Lys Asp 205 | 210 | 215 | 737 |
| aga ggc tgg gag acc tcc acc gac gct gtg atc aag tgg ctg agc gtc Arg Gly Trp Glu Thr Ser Thr Asp Ala Val Ile Lys Trp Leu Ser Val 220 | 225 | 230 | 785 |
| aac cgg gaa gga gtg gtt ttc ctg ttc tgg ggc tca tac gcc cat aag Asn Arg Glu Gly Val Val Phe Leu Phe Trp Gly Ser Tyr Ala His Lys 235 | 240 | 245 | 833 |
| aag gga gcg acc atc gac agg aaa cgt cac cat gtc ttg caa gct ctt Lys Gly Ala Thr Ile Asp Arg Lys Arg His His Val Leu Gln Ala Leu 250 | 255 | 260 | 881 |
| cat cca tct cct ttg tct gct cat cgt ggg ttc ctt ggt tgt aag cac His Pro Ser Pro Leu Ser Ala His Arg Gly Phe Leu Gly Cys Lys His 265 | 270 | 275 | 929 |
| ttc tcc aag gct aac ggg ctg ctg aaa cta tct ggg acg gag cct ata Phe Ser Lys Ala Asn Gly Leu Leu Lys Leu Ser Gly Thr Glu Pro Ile 285 | 290 | 295 | 977 |
| aac tgg aga gca ctc taactcttta tgctgcctta tactgttaac tgtttaaga Asn Trp Arg Ala Leu 300 | | | 1032 |
| tgaacatcac actatatttt ctacagcttt tccaagttca aaccaatcta taagcttca 1092 | | | |

tttgtctttt ggaatgatgc tgctttggc cggttttaga tacttaaaac actttaccac 1152
tctgccatgt tgactcatgt tcagtcaata taacttcac aacttgaaca aaaatgttat 1212
tttataattg attatattct gtacattaaa gattgtttt ttcccaggct gtttcataagg 1272
tactaggata ttaaactgtt attaacctat tttccatgat gtcaactgct taagtttta 1332
tgcatggaaaata aattatataat tta 1355

<210> 4
<211> 301
<212> PRT
<213> Gadus morhua

<400> 4
Met Ile Gly Gln Gln His Ile Asn Ser Phe Phe Ser Pro Val Ser Lys
1 5 10 15
Lys Arg Val Ser Lys Glu Leu Gly Lys Thr Glu Lys His Ala Glu Glu
20 25 30
Val Gln Ile Thr Pro Lys Lys Leu Arg Ser Ser Asn Val Glu Gln Lys
35 40 45
Thr Ser Ser Pro Gln Leu Ser Val Glu Gln Leu Glu Arg Met Ala Lys
50 55 60
Asn Lys Lys Ala Ala Leu Asp Lys Ile Arg Ala Lys Ala Thr Pro Ala
65 70 75 80
Gly Phe Gly Glu Thr Trp Arg Arg Glu Leu Ala Ala Glu Phe Glu Lys
85 90 95
Pro Tyr Phe Lys Gln Leu Met Ser Phe Val Ala Asp Glu Arg Ser Arg
100 105 110
His Thr Val Tyr Pro Pro Ala Asp Gln Val Tyr Ser Trp Thr Glu Met
115 120 125
Cys Asp Ile Gln Asp Val Lys Val Val Ile Leu Gly Gln Asp Pro Tyr
130 135 140
His Gly Pro Asn Gln Ala His Gly Leu Cys Phe Ser Val Gln Lys Pro
145 150 155 160
Val Pro Pro Pro Ser Leu Val Asn Ile Tyr Lys Glu Leu Cys Thr

| | 165 | 170 | 175 |
|---|-----|-----|-----|
| Asp Ile Asp Gly Phe Lys His Pro Gly His Gly Asp Leu Ser Gly Trp | | | |
| 180 | 185 | | 190 |
| Ala Asn Lys Gly Cys Cys Cys Leu Thr Arg Cys Leu Thr Val Arg Ala | | | |
| 195 | 200 | | 205 |
| His Gln Ala Asn Ser His Lys Asp Arg Gly Trp Glu Thr Ser Thr Asp | | | |
| 210 | 215 | | 220 |
| Ala Val Ile Lys Trp Leu Ser Val Asn Arg Glu Gly Val Val Phe Leu | | | |
| 225 | 230 | | 235 |
| Phe Trp Gly Ser Tyr Ala His Lys Lys Gly Ala Thr Ile Asp Arg Lys | | | |
| 245 | 250 | | 255 |
| Arg His His Val Leu Gln Ala Leu His Pro Ser Pro Leu Ser Ala His | | | |
| 260 | 265 | | 270 |
| Arg Gly Phe Leu Gly Cys Lys His Phe Ser Lys Ala Asn Gly Leu Leu | | | |
| 275 | 280 | | 285 |
| Lys Leu Ser Gly Thr Glu Pro Ile Asn Trp Arg Ala Leu | | | |
| 290 | 295 | | 300 |

<210> 5

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Artificial
Sequence - Primer used to generate cDNA of a
fragment of UNG gene

<400> 5

tacggctgcg agaagacgac agaagggg

27

<210> 6

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Artificial

Sequence - Primer used to prepare cDNA of a
fragment of UNG gene

<400> 6
tacggctccg agaagacgac agaa

24

<210> 7
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Artificial
Sequence - Primer used to generate cDNA portion
of cUNG gene

<400> 7
gghcargayc cctayca

17

<210> 8
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Artificial
Sequence - Primer used to generate cDNA portion of
gene

<400> 8
dccccasags agraavac

18

<210> 9
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Artificial
Sequence - Primer used to generate fragment of UNG
gene

<400> 9
tgtaccgaca ttgatggctt caagcat

27

```
<210> 10
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Artificial
      Sequence - Primer used to generate fragment of UNG
      gene

<400> 10
cccatccgct tagatctcca tgtccag                                27

<210> 11
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Artificial
      Sequence - Primer used to generate fragment of UNG
      gene

<400> 11
ccatccta atcgactcact atagggc                                27

<210> 12
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Artificial
      Sequence - Primer used to generate fragment of
      RACE fragment of UNG gene

<400> 12
atggaattcg attgagattg gcgccttgg                                30

<210> 13
<211> 30
<212> DNA
<213> Artificial Sequence
```

```

<220>
<223> Description of Artificial Sequence: Artificial
      Sequence - Primer used to construct rcUNG delta 74
      gene

<400> 13
accatggaat tcccaaaggc aacgcctgca                                30

<210> 14
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Artificial
      Sequence - Primer used to construct rcUNG delta 74
      gene

<400> 14
gagctcgatcg acttagagtg cctctccagt ttatagg                                37

<210> 15
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Artificial
      Sequence - Primer used to construct rcUNG delta 81
      gene

<400> 15
accatggaat tcttcggaga gacttggaga aga                                33

<210> 16
<211> 54
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Artificial
      Sequence - Primer used to construct rcUNG delta
      74o   gene

<400> 16

```

atggaattcg caaaagcaac gcctgcagg ttcggagaga cttggcgtcg tcag

54

<210> 17
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Artificial
Sequence - Primer used to construct rcUNG delta
81o gene

<400> 17
atggaattct tcggagagac ttggcgtcgt tgagctggct gc

42

<210> 18
<211> 47
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Artificial
Sequence - Primer used to prepare rcUNG gene

<400> 18
tctctcgaga aaagagaggc tgaagctccc attgacgatg aggatga

47

<210> 19
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Artificial
Sequence - Primer used to prepare rcUNG gene

<400> 19
gtagaattcg gatccatgtc tcctccagtc tagat

35